```
In [1]: import pandas as pd
        from sqlalchemy import create_engine
        from pandas import ExcelWriter
        from pandas import ExcelFile
```

Store XLSX into DataFrame

```
In [2]: xlsx_file = "Resources/us_wineries.xlsx"
        winery_data_df = pd.read_excel(xlsx_file, sheet_name='Sheet1')
        winery_data_df.head()
```

Out[2]:

	Winery_Name	State	Web_Site	Unnamed: 3
0	14 Hands	WA	www.14handswine.com	NaN
1	Abacela Vineyards & Winery	OR	www.abacela.com	NaN
2	Abarbanel Wine Co.	NY	www.kosher-wine.com	NaN
3	Abbott Winery	CA	www.abbottwinery.com	NaN
4	Abeja	WA	www.abeja.net	NaN

Clean DataFrame

```
In [3]: new_winery_data_df = winery_data_df[['Winery_Name', 'State', 'Web_Site']].copy
        ()
        new_winery_data_df.head()
```

Out[3]:

	Winery_Name	State	Web_Site
0	14 Hands	WA	www.14handswine.com
1	Abacela Vineyards & Winery	OR	www.abacela.com
2	Abarbanel Wine Co.	NY	www.kosher-wine.com
3	Abbott Winery	CA	www.abbottwinery.com
4	Abeja	WA	www.abeja.net

Store JSON data into a DataFrame

In [4]:
```python
json_file = "Resources/winemag-data-130k-v2.json"
wine_data_df = pd.read_json(json_file)
wine_data_df.head()
```

Out[4]:

	country	description	designation	points	price	province	region_1	region_2	taster_name
0	Italy	Aromas include tropical fruit, broom, brimston...	Vulkà Bianco	87	NaN	Sicily & Sardinia	Etna	None	Kerin O'Keefe
1	Portugal	This is ripe and fruity, a wine that is smooth...	Avidagos	87	15.0	Douro	None	None	Roger Voss
2	US	Tart and snappy, the flavors of lime flesh and...	None	87	14.0	Oregon	Willamette Valley	Willamette Valley	Paul Gregutt
3	US	Pineapple rind, lemon pith and orange blossom ...	Reserve Late Harvest	87	13.0	Michigan	Lake Michigan Shore	None	Alexander Peartree
4	US	Much like the regular bottling from 2012, this...	Vintner's Reserve Wild Child Block	87	65.0	Oregon	Willamette Valley	Willamette Valley	Paul Gregutt

Clean DataFrame

In [5]:
```python
# Create a filtered dataframe from specific columns
winetaster_cols = ["country","description", "designation", "points", "price",
"province", "region_1", "region_2", "taster_name", "taster_twitter_handle", "t
itle", "variety", "winery"]

winetaster_transformed= wine_data_df[winetaster_cols].copy()

# Rename the column headers
winetaster_transformed = winetaster_transformed.rename(columns={"country": "Co
untry",
                                                                "description":
"Description",
                                                                "designation":
"Designation",
                                                                "points": "Poin
ts",
                                                                "price": "Pric
e",
                                                                "province": "St
ate",
                                                                "region_1": "Re
gion_1",
                                                                "region_2": "Re
gion_2",
                                                                "taster_name":
"Taster_Name",
                                                                 "taster_twitte
r_handle": "Taster_Twitter",
                                                                "title": "Titl
e",
                                                                "variety": "Var
iety",
                                                                "winery": "Wine
ry"})

winetaster_transformed.head()
```

Out[5]:

	Country	Description	Designation	Points	Price	State	Region_1	Region_2	Taster_Name
0	Italy	Aromas include tropical fruit, broom, brimston...	Vulkà Bianco	87	NaN	Sicily & Sardinia	Etna	None	Kerin O'Keefe
1	Portugal	This is ripe and fruity, a wine that is smooth...	Avidagos	87	15.0	Douro	None	None	Roger Voss
2	US	Tart and snappy, the flavors of lime flesh and...	None	87	14.0	Oregon	Willamette Valley	Willamette Valley	Paul Gregut
3	US	Pineapple rind, lemon pith and orange blossom ...	Reserve Late Harvest	87	13.0	Michigan	Lake Michigan Shore	None	Alexande Peartree
4	US	Much like the regular bottling from 2012, this...	Vintner's Reserve Wild Child Block	87	65.0	Oregon	Willamette Valley	Willamette Valley	Paul Gregut

In [6]:
```python
#remove NaN
new_winetaster_transformed = winetaster_transformed[pd.notnull(winetaster_transformed['Price'])]
new_winetaster_transformed.head()
```

Out[6]:

	Country	Description	Designation	Points	Price	State	Region_1	Region_2	Taster_Name
1	Portugal	This is ripe and fruity, a wine that is smooth...	Avidagos	87	15.0	Douro	None	None	Roger Voss
2	US	Tart and snappy, the flavors of lime flesh and...	None	87	14.0	Oregon	Willamette Valley	Willamette Valley	Paul Gregut
3	US	Pineapple rind, lemon pith and orange blossom ...	Reserve Late Harvest	87	13.0	Michigan	Lake Michigan Shore	None	Alexande Peartree
4	US	Much like the regular bottling from 2012, this...	Vintner's Reserve Wild Child Block	87	65.0	Oregon	Willamette Valley	Willamette Valley	Paul Gregut
5	Spain	Blackberry and raspberry aromas show a typical...	Ars In Vitro	87	15.0	Northern Spain	Navarra	None	Michae Schachne

In [7]:
```python
new_winetaster_transformed.to_csv("Resources/clean_wine_list.csv", index=False, header=True)
```

In [8]:
```python
us_wine_list = new_winetaster_transformed.merge(new_winery_data_df, left_on='W
inery', right_on='Winery_Name', left_index=True)
us_wine_list.head()
```

Out[8]:

	Country	Description	Designation	Points	Price	State_x	Region_1	Region_2	Taster_Na
1742	US	A healthy addition of 13% Petite Sirah provide...	Classic	86	24.0	California	Napa Valley	Napa	Virg Bo
1742	US	Oaky and robust, this wine takes on a confecti...	None	86	17.0	California	Napa Valley	Napa	Virg Bo
1742	US	A generic expression of the variety with some ...	Classic Collection	84	19.0	California	Napa Valley	Napa	Virg Bo
1742	US	Made at an attractive price point, this wine d...	Classic Collection	90	24.0	California	Napa Valley	Napa	Virg Bo
1742	US	A great buy in a smooth, dry Merlot that shows...	None	90	22.0	California	Napa Valley	Napa	N

In [9]:
```python
clean_us_wine_list = us_wine_list.drop(columns=['Winery_Name', 'State_y'])
top_us_wines = clean_us_wine_list.rename(columns={"State_x": "State", "Web_Site": "Website"})
top_us_wines.head()
```

Out[9]:

	Country	Description	Designation	Points	Price	State	Region_1	Region_2	Taster_Na
1742	US	A healthy addition of 13% Petite Sirah provide...	Classic	86	24.0	California	Napa Valley	Napa	Virg Bo
1742	US	Oaky and robust, this wine takes on a confecti...	None	86	17.0	California	Napa Valley	Napa	Virg Bo
1742	US	A generic expression of the variety with some ...	Classic Collection	84	19.0	California	Napa Valley	Napa	Virg Bo
1742	US	Made at an attractive price point, this wine d...	Classic Collection	90	24.0	California	Napa Valley	Napa	Virg Bo
1742	US	A great buy in a smooth, dry Merlot that shows...	None	90	22.0	California	Napa Valley	Napa	N

Connect to local database

In [10]:
```python
import pymysql
pymysql.install_as_MySQLdb()

rds_connection_string = "root:password@127.0.0.1/top_wines_db"
engine = create_engine(f'mysql://{rds_connection_string}', pool_size=10, max_overflow=20)
```

Check for tables

In [11]: `engine.table_names()`

Out[11]: `['top_us_wine', 'wine_list', 'winery']`

Use pandas to load xlsx converted DataFrame into database

In [12]:
```python
new_winery_data_df.to_sql(name='winery', con=engine, if_exists='append', index=False)
```

Use pandas to load json converted DataFrame into database

In [13]:
```python
new_winetaster_transformed.to_sql(name='wine_list', con=engine, if_exists='append', index=False)
```

Use pandas to load new converted DataFrame into database

In [14]:
```python
top_us_wines.to_sql(name='top_us_wine', con=engine, if_exists='append', index=False)
```

Confirm data has been added by querying the winery table

- NOTE: can also check using pgAdmin

In [15]:
```python
pd.read_sql_query('select * from winery', con=engine).head()
```

Out[15]:

	winery_name	state	web_site
0	14 Hands	WA	www.14handswine.com
1	Abacela Vineyards & Winery	OR	www.abacela.com
2	Abarbanel Wine Co.	NY	www.kosher-wine.com
3	Abbott Winery	CA	www.abbottwinery.com
4	Abeja	WA	www.abeja.net

Confirm data has been added by querying the wine_list table

In [16]: `pd.read_sql_query('select * from wine_list', con=engine).head()`

Out[16]:

	country	description	designation	points	price	state	region_1	region_2	taster_name
0	Portugal	This is ripe and fruity, a wine that is smooth...	Avidagos	87	15.0	Douro	None	None	Roger Voss
1	US	Tart and snappy, the flavors of lime flesh and...	None	87	14.0	Oregon	Willamette Valley	Willamette Valley	Paul Gregutt
2	US	Pineapple rind, lemon pith and orange blossom ...	Reserve Late Harvest	87	13.0	Michigan	Lake Michigan Shore	None	Alexander Peartree
3	US	Much like the regular bottling from 2012, this...	Vintner's Reserve Wild Child Block	87	65.0	Oregon	Willamette Valley	Willamette Valley	Paul Gregutt
4	Spain	Blackberry and raspberry aromas show a typical...	Ars In Vitro	87	15.0	Northern Spain	Navarra	None	Michael Schachner

Confirm data has been added by querying the top_us_wine table

In [17]: `pd.read_sql_query('select * from top_us_wine', con=engine).head()`

Out[17]:

	Country	Description	Designation	Points	Price	State	Region_1	Region_2	Taster_Name
0	US	A healthy addition of 13% Petite Sirah provide...	Classic	86	24.0	California	Napa Valley	Napa	Virginie Boone
1	US	Oaky and robust, this wine takes on a confecti...	None	86	17.0	California	Napa Valley	Napa	Virginie Boone
2	US	A generic expression of the variety with some ...	Classic Collection	84	19.0	California	Napa Valley	Napa	Virginie Boone
3	US	Made at an attractive price point, this wine d...	Classic Collection	90	24.0	California	Napa Valley	Napa	Virginie Boone
4	US	A great buy in a smooth, dry Merlot that shows...	None	90	22.0	California	Napa Valley	Napa	None

In [ ]: